Final version of SYNT-D-25-01310,
for a special issue of Synthese on
Daniel Dennett: An Interdisciplinary
Perspective

DENNETT, NONHUMAN ANIMALS, AND CONSCIOUSNESS

David Rosenthal
Philosophy and Cognitive Science
Graduate Center, City University of New York
Email: davidrosenthal1@gmail.com
Mailing address: 425 Riverside Drive,
New York, NY 10025, USA
ORCiD: https://orcid.org/0000-0002-9323-0424

Abstract: Daniel Dennett's view about consciousness in nonhuman animals has two parts. One is a methodological injunction that we rely on our best theory of consciousness to settle that issue, a theory that must initially work for consciousness in humans. The other part is Dennett's application of his own theory of consciousness, developed in Consciousness Explained (1991), which leads him to conclude that nonhuman animals are likely never in conscious mental states. I defend the methodological injunction as both sound and important, and argue that the alternative approaches that dominate the literature are unworkable. But I also urge that Dennett's theory of consciousness and his arguments against conscious states in nonhuman animals face significant difficulties. Those difficulties are avoided by a higher-order-thought theory of consciousness, which is close to Dennett's theory, and provides leverage in assessing which kinds of mental state are likely to be conscious in nonhuman animals. Finally, I describe a promising experimental strategy for showing that conscious states do occur in some nonhuman animals, which fits comfortably with the higher-order-thought theory but not with Dennett's.

Keywords: Daniel Dennett, animal consciousness, behavioral markers, first-person operationalism, blindsight, change blindness, higher-order-thought theory

Data availability statement: not applicable

Prologue: Daniel Dennett and I first talked extensively about consciousness in the spring of 1990 at the Center for Interdisciplinary Studies (ZiF) in Bielefeld. That led to several published discussions of one another's views. In his last published comment on my work, Dennett raised a challenge about higher-order theories in connection with the conscious states of nonhuman animals. This essay, to my sadness too late for him to see, is my reply.

I. Dennett's Methodological Injunctions

Daniel Dennett's penetrating work about consciousness has had a deep impact, inspiring a great range of commentary, both admiring and critical. There is likely no body of work that has had a greater influence on the literature about consciousness. My own conviction is that Dennett's theory of consciousness is very much along the right lines, though I will also argue that his theory faces some difficulties, which require adjustment.

But despite extensive discussion of Dennett's writing on consciousness, there is one aspect of that work that has received very little attention. That is Dennett's view about consciousness in nonhuman animals. That may in part be due to the brevity of Dennett's own explicit discussion of that issue, which occurs late in Consciousness Explained (1991, pp. 442-449), in somewhat greater detail in a subsequent article (1995), and occasionally elsewhere somewhat in passing (e.g., 2018, 168f., and 2022).

Still, that topic has recently attracted significant attention. In what follows, I describe and assess Dennett's view about consciousness in nonhuman animals, comparing it to other recent writing about that issue.

Dennett's view about consciousness in nonhuman animals is best seen as having two parts. One part consists of a methodological prescription about how the issue must be addressed. We must, he writes, "first devis[e] a theory that concentrate[s] exclusively on human consciousness … and then look and see which features of that account apply to which animals, and why" (1995, 700; Dennett's emphasis).

This methodological injunction is itself best seen as having two parts. We must apply our best theory of consciousness to the case of nonhuman animals, and because we understand what consciousness is from its occurrence in humans, any sound theory must be grounded in human consciousness. The other part of Dennett's view, then, is his application of his own theory to nonhuman animals.

I'll first argue at some length that Dennett's methodological considerations are both sound and important. In the second half of my discussion, then, I'll argue that Dennett's theory of consciousness distorts the issue about nonhuman animals, and leads to conclusions that are unfounded and likely erroneous. The current literature urges conclusions that differ starkly from Dennett's, but I'll argue that it also distorts the issue in much the same way as Dennett's theory does. Establishing whether conscious states occur in nonhuman animals is, I'll urge, considerably more difficult than either Dennett or the dominant literature envisage.

Dennett's methodological injunctions have conspicuously not found favor in the current literature about consciousness in nonhuman animals. But the objections to those injunctions have not been compelling. Jonathan Birch rejects Dennett's insistence that we appeal to a theory of consciousness (e.g., 2022, 134). But his rejection largely relies on his dismissing current theories because there is no consensus about which is correct. And lack of consensus cannot by itself show a theory to be wrong, any more than consensus by itself can show that a theory is right. One must always evaluate the merits of specific theories.

Kristin Andrews (2024) also rejects appeals to current theories of consciousness, appealing to Colin Allen and Michael Trestman's (2016) contention that no current theory is sufficiently secure to accept. But Allen and Trestman also rely on lack of consensus, rather than assessing specific theories in a serious way. Similarly, Tim Bayne and Nicholas Shea also reject reliance on any current theory, since no current theory "enjoys widespread support" (2020, 69). And Andrews, Birch, and Jeff Sebo (2025, p. 822) note that there are 22 current theories of consciousness, as though to suggest current theorizing about consciousness is not to be taken seriously. But most by far of those 22 are widely seen as fringe at best, with only a very few being seen as real contenders, not too many to take seriously and evaluate on their merits.

Bayne and Shea also urge that the measures of consciousness used to test a theory may derive from that theory, thereby undermining the reliability of such a test (2020, 69). But they offer no reason to think that there cannot be reliable measures that are independent of the theories being tested.

In any case, even if no current theory were acceptable, that would not affect Dennett's methodological injunction that we rely on our best theory. We would then need to develop a theory that is acceptable. Prior to modern mathematical physics there was no acceptable theory of motion, but that was no reason to deny that we must understand physical motion only by appeal to our best theory. It is patently unscientific to reject the need for a sound theory.

Andrews (2024) also rejects Dennett's injunction that our best theory must be anchored in the case of humans. She maintains that a secure theory must adopt a default assumption that consciousness does occur in nonhuman animals generally. Such a theory would then simply address the different ways in which consciousness occurs in various species (see also Andrews 2016). In a somewhat similar spirit, Birch (e.g., 2022) rejects theories that rely on the human case because the conditions for consciousness in humans would have to be relaxed to apply to nonhuman animals.

Both contentions simply ignore Dennett's point that we understand what the property of consciousness itself is by appeal to the human case. Without that anchor, which is needed to fix the reference of the term, 'conscious', we would be left to apply the term to other species impressionistically, and so to a variety of different mental properties that might have little if anything to do with mental states' occurring consciously.

Andrews, Birch, and Sebo (2025) complain that the lack of an accepted theory of consciousness means that there is no accepted definition of what consciousness is. But strict definitions aside, it is plain that there is no way to fix what our notion of consciousness applies to except by appeal to the human case.

The foregoing objections to one side, there are highly compelling reasons to endorse both of Dennett's methodological injunctions. The only serious alternative to applying our best theory of consciousness would be to rely on behavioral markers that appear in a commonsense way to indicate consciousness (Allen & Trestman 2020), a strategy endorsed in almost all of the current literature (e.g., Allen & Trestman 2016; Andrews 2020, ch. 5, and 2024; Andrews et al, 2025; Bayne & Shea 2020; Birch 2022 and 2024; Birch et al 2020; Ginsburg & Jablonka 2019; Godfrey-Smith 2020; Veit 2023).

Such markers can indicate the occurrence of mental states with reasonable reliability. Indeed, their use with nonhuman animals would be a nice application of Dennett's well-known intentional stance, on which we impute psychological states as the best explanation of rational behavior (e.g., Dennett 1971, 1989). Relying on such markers also to determine whether nonhuman animals are in mental states that are conscious would make that assessment relatively easy and straightforward, resulting in many mental states for many nonhuman animals.

But though behavioral markers are arguably good for revealing the occurrence of mental states, mental states are not always conscious states. So it could very well be that the mental states that behavioral markers often reveal are simply unconscious mental states.

It is evident from the human case that mental states are not always conscious. Perception is sometimes subliminal, and many thoughts and desires occur without being conscious. Behavioral markers have no way to differentiate mental states that are conscious from those that are not. Such markers can at best differentiate states that are mental from those that are not. Distinguishing conscious from unconscious mental states would require some theoretical understanding of how the conscious cases differ from those that occur unconsciously. We would need to apply our best theory of consciousness, as Dennett enjoins.

Though advocates of behavioral markers occasionally note this challenge (e.g., Birch 2022; Andrews et al 2025), they tend to downplay or disregard it. It is typically assumed that mental states occur consciously by default or at least very often (e.g., Andrews 2024). Absent that assumption, reliance on behavioral markers could not be sound. But behavioral markers themselves do nothing to support an assumption of default or frequent conscious occurrence, and no serious independent support is offered. So that assumption seems to be a question-begging stipulation that consciousness does occur in nonhuman animals.

Indeed, absent any independent support, it is tempting to see that assumption as in effect a solution to the problem of other minds for nonhuman animals. But that would seriously misconceive the issue. The problem of other minds is how to ward off skepticism about whether people other than oneself are in conscious mental states. The question about nonhuman animals is not about resisting philosophical skepticism. It is a factual and scientific question about whether they, like humans, are in mental states that are conscious, a question that invites applying relevant theoretical considerations.

Dennett himself actually denies that unconscious states are, strictly speaking, mental states, for reasons to be reviewed in section IV. That unusual denial will be important for assessing Dennett's theory of consciousness, but it is not relevant to his methodological injunctions.

An advocate of behavioral markers might insist that some markers are so compelling that it would be difficult to believe that the relevant behavior is not due to conscious mental states, perhaps because similar behavior in humans typically results from conscious states (e.g., Andrews et al, 2025). That would be a failure of imagination.

The type of behavior a mental state elicits is due to that state's representational content. Behavior reveals the representational content of mental states, not whether those states are conscious. So states that occur consciously in humans might occur unconsciously in nonhuman animals.

It is sometimes held that behavioral markers should be supplemented by neural markers. But neural markers are highly unlikely to help. For one thing, as Birch notes, neural anatomy and function vary significantly from one species to another (2022, 145). And we would in any case need a sound theory of consciousness to show how neural markers could differentiate conscious from unconscious mental states.

There is sound reason, then, to adopt Dennett's precept that we rely on our best theory of consciousness. What about his second precept, namely, that a sound theory must rely on what we know about consciousness in the human case?

As already noted, our understanding of what consciousness is rests firmly on the human case. We fix what our notion of consciousness refers to in that way. Without such an appeal, there would be no way to prevent the term, 'conscious', from being applied in ways that are inescapably impressionistic.

Indeed, it is on its face odd to think we could address whether consciousness occurs in nonhuman animals without relying on our understanding of consciousness in humans. We must always in investigating any phenomenon give priority to the cases we understand best, leaving it open that we may come to revise what we initially thought about those cases.

We rely on sincere verbal report to determine whether mental states are conscious in the human case. Verbal reports are the gold standard for determining whether a mental state occurs consciously, since they express the first-person access one has to one's own conscious states. And because verbal report is unavailable in nonhuman animals, one might infer that the human case is simply irrelevant for nonhuman animals.

That would be a mistake. Determining whether mental states are conscious will plainly be significantly more difficult with nonhuman animals than with humans. But that is something we must simply acknowledge and deal with. And it is likely that we can make progress only by appeal to a theoretical understanding of how conscious mental states differ from mental states that are not conscious. More in sections V-VII.

The next two sections continue to give support for Dennett's methodological injunctions. Section II reinforces the difficulties for behavioral markers, which are the only serious alternative to applying our best theory. And because behavioral markers cannot show that the mental states they detect are conscious, section III considers independent reasons that have been offered for holding that all mental states are conscious.

Section IV, then, turns to Dennett's theory of consciousness and its application to nonhuman animals. After briefly sketching the theory, I argue that it faces difficulties, and is actually in conflict with other important aspects of Dennett's thinking. I also argue that the higher-order-thought (HOT) theory of consciousness (Rosenthal 1993, 2005, 2022, 2025b) is a close cousin of Dennett's theory, and that it avoids the difficulties that face his theory.

Sections V and VI argue that the HOT theory also sheds light on which types of mental state are likely to occur consciously in nonhuman animals, an advantage for the HOT theory over Dennett's. And section VII concludes by describing an experimental strategy for establishing that some mental states in nonhuman animals are conscious. It also explains why that strategy does not fit comfortably with either Dennett's theory or much of the current literature, though it does fit naturally with the HOT theory.

II. More on Markers

There is a way of construing the question about consciousness in nonhuman animals that might seem to invite the appeal to behavioral markers. If the question is whether animals themselves are conscious, that might amount simply to whether they are capable of being awake and receptive to sensory input. That is what Andrews calls awake consciousness (2020, 75) and I have elsewhere called creature consciousness (e.g., Rosenthal 1993).

And behavioral markers are useful for determining whether nonhuman animals alternate between sleep and waking conditions and whether when awake they are responsive to sensory input. And there can be no serious doubt that all but relatively primitive forms of animal life are conscious in that way. Insects, for example, alternate between sleep and waking states, and when awake they are responsive to sensory inputs (Helfrich-Förster 2018).

That is not the question that concerns most people. Rather, we want to know whether nonhuman animals are in mental states that are conscious, what I have elsewhere called state consciousness. And behavioral markers cannot by themselves distinguish conscious from unconscious mental states. Still, being in mental states, whether conscious or not, likely suffices for nonhuman animals to be responsive when awake to sensory inputs. And behavioral markers can reveal whether they are in mental states. So lack of clarity about what the question about nonhuman animals is can seem to invite appeal to behavioral markers.

Andrews is clear that the issue is whether the mental states of nonhuman animals occur consciously. And perhaps because she recognizes that behavioral markers cannot by themselves distinguish conscious from unconscious mental states, she tends to dismiss cases of unconscious mental states altogether. She takes unconscious driving to be paradigmatic of unconscious sentience, and correctly notes that such cases are actually just inattentive, rather than genuinely unconscious (2020, 83). And if alleged unconscious perception were always just of lack of attention, perception would never be unconscious, strictly speaking. Behavioral markers that reveal perceptual functioning would then perforce reveal conscious perceiving.

But extensive experimental findings with humans make it overwhelmingly clear that sentient states do not always occur consciously. Masking studies, for example, provide conclusive evidence of unconscious perceptual states in humans (e.g., Marcel 1983; Öğmen and Breitmeyer 2006). Unconscious change detection also plainly occurs (Fernandez-Duque & Thornton 2000, Thornton & Fernandez-Duque 2001; replicated by

Laloyaux et al 2006), a phenomenon that will figure in the experimental strategy proposed in section VII for showing that conscious states do occur in some nonhuman animals. And sentience aside, there is also powerful evidence of unconscious thinking and volition (e.g., Dijksterhuis et al 2006, Waroquier et al 2010, and Sklar et al 2012; Peirce & Jastrow 1884).

These experimental findings are reflected in everyday experience. People sometimes have desires or thoughts that they are wholly unaware of, and indeed firmly deny having. And people occasionally respond to visual stimuli without being in any way aware of having seen them. We tend to disregard such cases, though circumstances do occasionally compel us to take account of them.

Experimental work shows decisively that states occur in humans that, though not conscious, have behavioral effects similar to those of conscious mental states. So when similar behavioral effects occur in nonhuman animals, they may well result simply from unconscious states. Since specific markers can occur with states that are not conscious, it is hard to see how behavioral markers of states' being conscious could get serious traction.

Birch (2002) advances what he calls the facilitation hypothesis, on which conscious states facilitate behavior not facilitated by comparable unconscious states. Birch describes his appeal to the facilitation hypothesis as "theory-light," apparently because it is compatible with a number of otherwise disparate theories of consciousness. But the facilitation hypothesis is simply an empirical claim offered in support of behavioral markers, and is in no way theoretical.

And the facilitation hypothesis itself is highly questionable. The relevant unconscious states are typically significantly weaker in representational signal strength than comparable conscious states, notorious with masked visual states (Öğmen & Breitmeyer 2006), but true in general. Support for the facilitation hypothesis would require comparable conscious and unconscious states that are roughly equivalent in representational signal strength, since otherwise the difference in signal strength would be the best explanation for a difference in facilitation.

An advocate of behavioral markers might insist that, though mental states do sometimes occur unconsciously, occurring consciously is nonetheless the default. So we can assume that they are conscious absent some reason to the contrary. Behavioral markers would then reveal conscious mental states much of the time.

But it is by no means obvious that conscious occurrence is the default even in humans. As just noted, we tend in everyday life to pay no attention to or even acknowledge mental states that might occur unconsciously. But that by itself is no reason to think that it is unusual for them to do so. And even if conscious occurrence were the default in humans, we would need some independent reason to think that it also is with nonhuman animals.

Behavioral markers are the only serious alternative to applying a theory, and Dennett urges that any sound theory must be anchored in the human case. So it is worth stressing that behavioral markers themselves are perforce also anchored in the case of

human consciousness, imputing mental states to nonhuman animals on the basis of behavior that would indicate mental states in humans. Reliance on the human case is unavoidable.

Bayne and Shea (2020) have advocated a variant of the marker approach that accommodates some appeal to theoretical considerations. They take consciousness to be a natural kind, and urge reliance on "a variety of dispositions and capacities ('marks') that are putatively associated with consciousness." These marks can include "various pre-theoretic indicators of consciousness," but also those that depend on theoretical considerations (2020, 70). Bayne and Shea advocate appeal to marks that cluster in the way characteristic of a natural kind (2020, 71). But such clusters cannot help unless the individual marks are themselves reliable. And unless we evaluate theoretical considerations by appeal to an overall theory of consciousness, those considerations may well be as impressionistic as the marks themselves.

It is sometimes held that we must understand consciousness in terms of the benefit it provides the organism (e.g., Humphrey 2002; Dennett 2022). And if there were significant utility for a state's being conscious, perhaps when behavioral markers do reveal a mental state, that state will often be conscious (e.g., Veit 2023; Ginsburg & Jablonka 2019).

But the appeal to utility is highly questionable. The utility a mental state has for the organism is due primarily to the state's representational content (Rosenthal 2008). It is unclear how any mental state could have utility except by way of such content, and it would be rare for a state's being conscious to add to that content. And because it takes longer for perceptual input to become conscious than for it to affect behavior (Fleming & Michel 2024); our intuitive sense that a state's being conscious typically figures in determining behavior is likely illusory. So it is unlikely that a state's being conscious adds significant utility if any at all to the utility that results from that state's representational content (Rosenthal 2008; Michel, manuscript; Robinson et al 2015). This undermines Birch's claim that consciousness facilitates "a cluster of cognitive abilities" (2022, p. 140; see also Gottlieb 2024, 86-7).

The idea that mental states occur consciously by default may seem tempting because of Thomas Nagel's contention that sentience can occur only in connection with a conscious point of view (e.g., 1974). But perception in nonhuman animals is demonstrated by the detection and discrimination of stimuli, which cannot by itself establish any such conscious perspective. Indeed, it is arguable that we should understand sentience generally by appeal to the discriminative role of perceiving, rather than in terms of conscious subjectivity (Rosenthal 2010, 2022, 2025a).

Much of the experimental evidence of unconscious states in humans does not involve any elaborate, sustained behavior. So advocates of behavioral markers might urge that when behavior in nonhuman animals is elaborate and sustained, we can safely infer that it does result from mental states that are conscious (Andrews et al, 2025, pp. 823-824). But unconscious states in nonhuman animals do sometimes result in elaborate, sustained behavior that is altogether indistinguishable from behavior that characteristically results from conscious mental states.

When some part of primary visual cortex (V1) is destroyed in humans, the person can no longer consciously see readily visible stimuli presented in the affected part of the visual field. Nonetheless, forced-choice guessing about the nature of the stimuli are strikingly accurate, typically in excess of 80% (Weiskrantz 1990, 1997; Cowey 2010). Such accuracy plainly demonstrates that relevant stimulus properties are visually registered, though not consciously. Such unconscious vision can even guide action (Danckert and Rossetti 2005; Stoerig 2011). This phenomenon, known as blindsight, is a dramatic case of unconscious perception.

Blindsight has also been studied in nonhuman animals. In Lawrence Weiskrantz's Cambridge University lab in the 1960s a rhesus macaque monkey named Helen had V1 surgically ablated. After a period of adjustment she was able accurately and spontaneously to use visual input to navigate around physical objects, pick up tiny objects, and even catch flies in midair (Humphrey & Weiskrantz 1967; Humphrey 1974; Cowey & Stoerig 1995; Weiskrantz 1997, ch. 6). An online video of Helen's behavior (Humphrey 1972) is striking.

We know from humans that V1 lesions result in vision no longer being conscious in the affected area. So it is beyond serious question that Helen's surgery resulted in total loss of conscious vision. Still, after the adjustment period, Helen regained visual functioning that was spontaneous, sustained, and highly focused. So unconscious visual states plainly can produce such behavior. The representational signal strength of Helen's unconscious visual states facilitates behavior, undercutting Birch's facilitation hypothesis. Elaborate, sustained perceptual functioning cannot by itself demonstrate conscious perceptual states in nonhuman animals. And Dennett's injunction that we anchor our investigation in the human case is again pivotal.

There is an esoteric worry about Helen. It has recently been found that if V1 damage in monkeys occurs in the first year of life, the pulvinar will take over lost V1 functions (Bourne & Morrone 2017; my thanks to Kathleen Akins for alerting me to this). And then the resulting vision might well be conscious. But that finding is not relevant to Helen, who was at least two years old when V1 was ablated (Humphrey 1974, and personal communication).

There is only one known human case of total blindsight like Helen's, a patient known as TN. TN can also visually avoid obstacles, but unlike Helen cannot do so fluently or comfortably (de Gelder et al 2008; for complications see de Gelder et al 2024). One might speculate that because TN's V1 damage occurred at the age of 52, life-long habituation to rely on conscious vision may have resulted in that lack of fluency. Putting TN's lack of fluency aside, Helen's case is sufficient to show that unconscious vision can result in fluent, elaborate behavior in nonhuman animals. (See Michel, "9," in Birch et al 2022, for somewhat similar considerations.)

The human case makes it clear that none of Helen's post-surgery visual states were conscious. But such surgery is a strikingly unusual circumstance. So one might urge that Helen's case does nothing to undermine the assumption that the default for mental states in nonhuman animals is that they are conscious.

But that would be too quick. We know that Helen's post-surgery vision was unconscious. But we cannot take for granted that her pre-surgery vision was conscious; that would need to be established empirically. Indeed, if her pre-surgery vision was unconscious, she might post-surgery simply have regained her pre-surgical unconscious vision, which would explain the remarkable fluency of her post-surgery vision-based behavior. So Helen's case does not by itself support the claim that the mental states of nonhuman animals are conscious by default. Still, we will see evidence in section VII that her pre-surgery vision was almost certainly conscious.

More generally, we need serious experimental techniques to show that any mental states of nonhuman animals are conscious. Those techniques must be able to distinguish conscious mental states from mental states that are not conscious. Behavioral markers cannot do that, since they can occur with both conscious and unconscious states. Section VII will describe an experimental strategy for drawing that distinction, one that comports with Dennett's methodological injunctions.

III. Denials of Unconscious Mentality

The only alternative to reliance on our best theory of consciousness is some version of behavioral markers. But unconscious states in nonhuman animals can result in behavior altogether indistinguishable from that produced by conscious mental states. And there is no independent support for holding that mental states by default occur consciously in nonhuman animals. The appeal to behavioral markers arguably resembles optimistic guesswork, perhaps based on the belief that nothing better is possible, though a sound theory would provide appropriate scientific tools.

But behavioral markers are reliable in detecting mental states, independently of their being conscious. And some have claimed that perceptual states are invariably conscious, even those in blindsight and in masked priming and similar experimental techniques. The mental states that behavioral markers reveal in nonhuman animals would then perforce be conscious, and we would not need to rely on a theory. So we must examine the reasons offered for holding that very strong claim.

One reason stems from a debate about how to determine whether perceptual states are conscious (Dienes 2008; Timmermans & Cleeremans 2015). On the so-called objective measure, a perceptual state is conscious if detection of perceptual information is accurate above chance. The subjective measure, by contrast, relies on verbal report or its equivalent to determine whether a perceptual state is conscious.

It might seem that the subjective measure is simply inapplicable to nonhuman animals, since they cannot report being in mental states. And as will emerge in section IV, Dennett raises serious difficulties for behavior in nonhuman animals that might be inaccessible to verbal report. So some might insist on using an objective measure for nonhuman animals.

And the objective measure fits comfortably with reliance on behavioral markers. On the objective measure, a perceptual state is conscious if it is accurate above chance. But a state that is not accurate would not count as perceptual to begin with. So the

objective measure counts all perceptual states as conscious, even if the individual firmly denies perceiving anything (e.g., Phillips 2018). On the objective measure, any perception revealed by a marker is perforce conscious.

But this result departs dramatically from our commonsense conception of what it is for a state to be conscious. If one is wholly unaware of a mental state, that state is not conscious; a state is conscious only if one is in some way aware of being in that state. And the verbal report that the subjective measure relies on expresses one's awareness of being in that state. It is at best a degenerate conception of consciousness on which a state counts as conscious even if one firmly denies being in it.

Some researchers opt for an objective measure because the introspective access required for a subjective measure can be unreliable (cited by Timmerman & Cleeremans 2015, 21-22). That consideration is also unconvincing. Introspection is not flawless, but no empirical measure ever is. And there is no reason to doubt that introspective access is typically accurate. We cannot rely on the objective measure to sustain an appeal to behavioral markers.

There is a somewhat more sophisticated consideration that has been offered in support of holding that all perceptual states are conscious. Signal-detection theory posits a criterion that individuals operate with in determining whether to report being in a mental state. If that criterion is set conservatively, a state must be more intense for an individual to report being in that state, whereas if the criterion is set liberally, less intensity is required for the individual to report the state. Whether one is disposed to report a state depends on how the signal-detection criterion is set.

The apparent problem is that how the criterion is set itself varies with factors such as attention and alertness, which may seem irrelevant to whether a state is conscious. So it has been argued (e.g., Phillips 2018) that verbal report cannot determine whether a state is conscious. And that would leave only the objective measure.

This is a mistake. The tendency to report a state goes hand in hand with that state's being conscious. The reason a conservative criterion requires more intensity for an individual to report being in that state is that the state must be more intense for the individual to be aware of it. And a state's being conscious consists in one's being subjectively aware of it.

So on a conservative criterion a state must be more intense for that state to be conscious. Similarly, less intensity is required for a state to be conscious on a liberal criterion. Variation of the criterion does not result in reports being inaccurate about which states are conscious. The criterion determines both whether one is disposed to report and whether the state is conscious.

Indeed, the argument from variability of the criterion simply assumes some unstated condition for a state to be conscious, a condition that is independent of verbal report, thereby begging the question against the subjective measure of consciousness. Criterial variability does not support the objective measure, and cannot sustain the denial of unconscious mental states.

Indeed, the labels for the two measures of consciousness arguably misrepresent the situation. We must determine objectively whether a state is conscious. But consciousness is a matter of how one's mental life subjectively appears to one. So that is what we must objectively determine. And the so-called subjective measure is an objective way to do that. The problem is how to devise an objective procedure to test that in nonhuman animals. As noted earlier, section VII proposes a theoretically sound experimental strategy for doing so.

IV. Dennett's Theory of Consciousness

Behavioral markers cannot distinguish conscious from unconscious mental states. The only alternative is to apply our best theory of consciousness to the case of nonhuman animals. And because we must rely on the human case simply to understand what it is for a mental state to be conscious, that theory must be anchored in consciousness in humans. We have strong support for Dennett's methodological injunctions.

Dennett's own theory of consciousness does rely on the human case. And there is a very great deal to recommend that theory. But there is also reason to question whether it is the best theory.

Dennett introduces his theory by way of some puzzles about timing (1991, chs. 5-6; Dennett & Kinsbourne 1992). One case involves seeing a long-haired woman without eyeglasses, though that perception is immediately contaminated by a memory of a different woman, short-haired and with eyeglasses. The result is a conscious visual experience of a short-haired woman with eyeglasses (1991, 117-8). Memories do sometimes skew conscious perception in this way.

The puzzle Dennett poses is this. The contaminating memory intrudes after retinal stimulation but before conscious experience. But when? Does the visual input of a long-haired woman make it to consciousness and then get replaced by a conscious visual experience of a short-haired woman? Or does the memory alter the perceptual content before it becomes conscious?

Dennett contends that there is no fact of the matter about when the memory intrudes. Nothing in our subjective experience or in our folk-psychological picture of things can decide between them. We can describe things either way, but Dennett claims that nothing in reality can decide between the two alternatives.

Dennett also presents several experimental findings that seem to pose the same puzzle. One involves the so-called phi phenomenon, in which two lights at different locations briefly flash in succession. If the spatial and temporal distances are just right, observers don't experience two distinct flashes, but rather one light moving back and forth between the two locations (Dennett 1991, 114; Kolers & von Grünau 1976).

Retinal input here is plainly of two lights flashing. And if the second flash never occurred, one would consciously experience just the first light flashing. So the second flash somehow alters what one consciously experiences. The puzzle: Does input from that first flash make it to consciousness and then get replaced by a conscious

experience of the light moving back and forth? Or does that replacement occur before the first flash becomes conscious? Dennett again urges that since nothing in subjective experience or folk psychology could decide which happens, there is nothing in reality that could do so.

Dennett concludes that, since nothing in reality could decide between these alternatives, "there is no reality of consciousness independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)" (1991, 132). This is Dennett's first-person operationalism (FPO), according to which "there are no fixed facts about the stream of consciousness independent of particular probes" (1991, 138; cf. 275).

As Michael A. Cohen and Dennett note, the effects that reveal that some perceptual input has made it to consciousness "are all the products of cognitive access" (Cohen & Dennett 2012, 140), mainly "verbal report, button pressing etc." (Cohen & Dennett 2011, 358). For that reason, FPO "brusquely denies the possibility in principle of consciousness of a stimulus in the absence of the subject's belief in that consciousness" (Dennett 1991, 132). So FPO unequivocally endorses the subjective measure of consciousness, and does so in terms that echo a higher-order-thought theory of consciousness. More on that shortly.

One might urge that knowing the neural correlates of the relevant psychological events could settle when the revision of content occurs. But isolating the relevant neural events would require knowing which neural events correlate with the revision of content and with the onset of consciousness. And we cannot establish such correlations without some independent way to determine the timing of those psychological events, themselves.

It is common for perceptual inputs to compete to become conscious. But Dennett holds that prior to becoming conscious each of the competing states is nothing more than a draft of a potential conscious state; hence his description of his view as the multiple-drafts model of consciousness (1991, 111ff.). And because competing states prior to consciousness are mere drafts, Dennett also contends that they do not qualify as mental occurrences, properly so-called. He regards such drafts as mere "events of content-fixation" (e.g., 1991, 365), what he had earlier called subpersonal states (1969, 95).

So Dennett counts no state as genuinely mental unless it is conscious. Despite endorsing a subjective measure of consciousness, Dennett shares with the objective measure its striking denial that any unconscious states are, properly speaking, mental.

Still, it is important that Dennett arrives at that denial in a way different from the objective measure. The objective measure counts a perceptual state as conscious only if it is reasonably accurate perceptually. Since no state counts as perception if it is not accurate to some degree, being a perception coincides with being conscious.

Dennett does not use perceptual accuracy as a measure of consciousness. Rather, he holds that a state is conscious only if it has suitable effects, primarily though not exclusively verbal report or the equivalent. And because unconscious states are mere

drafts of mental states, they are competing to have such effects, but do not yet have any. So they are not properly speaking mental.

This difference between Dennett and the objective measure of consciousness is crucial in connection with nonhuman animals. On the objective measure all perceptual states are conscious. So any nonhuman animal that functions perceptually will, on an objective measure, be in many states that are conscious. But the perceptual states of nonhuman animals do not result in the effects that Dennett maintains are necessary for a state to be conscious. So he concludes not only that they fail to be conscious, but also that because they are mere drafts, they are not even mental, properly speaking.

Verbal report is indeed unavailable in nonhuman animals. And Dennett argues that nonverbal substitutes in nonhuman animals are unreliable. Because button presses and other nonverbal substitutes require extensive training, which relies on reinforcement conditioning. So for all we know, those substitutes may simply reflect a conditioned response to states that are not conscious (2018, 168). There is, Dennett concludes, no reliable way to detect conscious states in nonhuman animals.

Dennett's concern about reinforcement conditioning is sound. Indeed, it is similar to the foregoing concerns about behavioral markers, which also cannot distinguish conscious from unconscious mental states. So directly testing whether some state in a nonhuman animal is conscious is likely not the way to go. But an indirect test of whether particular states in nonhuman animals fail to be conscious can apply despite the concern about reinforcement conditioning. Such tests will be described in section VII, and hold promise for showing that some mental states in nonhuman animals are conscious. Dennett's concerns are not the last word.

Advocates of behavioral markers urge that mental states by default are conscious, in nonhuman animals as well as humans. And Dennett contends that all mental states are conscious. But that contention does not support appeal to markers, since the states they detect in nonhuman animals are not even mental, properly speaking. We cannot avoid appeal to our best theory.

Putting aside for now the issue about reinforcement conditioning, there are reasons to question whether Dennett's theory of consciousness is sound. One reason stems from Dennett's contention that "there is no reality of consciousness independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)" (1991, 132).

The effects Dennett appeals to are pivotal for detecting the occurrence of conscious states in humans. But factors that enable detection do not always reveal the nature of the phenomenon being detected. And it is plain from first-person access that the overwhelming majority if conscious states occur with no behavioral effects. Moreover, most conscious states are so fleeting that they do not even affect memory.

Why, then, are the effects Dennett cites useful for detection? Principal among those effects are verbal reports that one is in a particular mental state. Such reports are useful because they express an awareness of the state that is reported. A speech act must express a thought with the same content as that of the speech act. So if I say

that I see something, I express my thought that I do, thereby revealing my awareness that I do.

Verbal reports are useful for detecting conscious states because they express an awareness of that state, in virtue of which the state is conscious. But that awareness need not be expressed for the state to be conscious. The effects Dennett appeals to are not necessary for a state to be conscious, though they are typically sufficient. What is necessary is that an individual is aware of being in the state, and verbal report is evidence of such awareness.

This foregoing reasoning is the basis of the higher-order-thought (HOT) theory of consciousness that I have developed elsewhere (Rosenthal 1993, 2005, 2022, 2025b). Since no state of which one is wholly unaware is conscious, a necessary condition for a state to be conscious is that one is aware of it. The behavioral effects that Dennett appeals to actually point to holding that a state is conscious only if one has a HOT that one is in that state.

Indeed, as noted earlier, Dennett in effect endorses the main plank of the HOT theory when he rejects the "possibility in principle of consciousness of a stimulus in the absence of the subject's belief in that consciousness" (Dennett 1991, 132). And Dennett has independently noted that he sees merit in the theory (e.g., Dennett 1991, 307-309; 2018; 2022). So we can reasonably regard the HOT theory as a close cousin of Dennett's own theory of consciousness (Rosenthal 2018, §IV). But since the appeal to behavioral effects is better seen as a tool for detecting conscious states than as a necessary condition for a state to be conscious, there is reason to prefer the HOT theory over Dennett's own theory.

The other reason to question Dennett's theory concerns his contention that unconscious states are mere drafts of mental states, and so do not qualify as genuinely mental. This surprising claim not only fails to square with our folk-theoretic views; it is also actually in conflict with other important views long held by Dennett.

As noted earlier, Dennett endorses using what he calls the intentional stance to explain psychological behavior (1971, 1989). On the intentional stance, we ascribe mental states to others to make sense of their behavior on an assumption of overall rationality. But behavior is often rational when it is due to mental states that, by any commonsense standard, fail to be conscious. You might use the intentional stance to ascribe to me a belief or desire even if I am unaware of having that belief or desire, and sincerely deny that I do.

So the intentional stance underwrites the ascription of unconscious states that are genuinely mental. Such unconscious states often have ongoing psychological roles, and so often cannot be mere drafts. The intentional stance undermines Dennett's denial of mental status to unconscious states (Rosenthal 2018), as Dennett himself has acknowledged (2018).

Because a mental state is conscious on the HOT theory only if one has a HOT that one is in that state, countenancing unconscious states as genuinely mental fits comfortably with the theory. That again favors the HOT theory over Dennett's. And the appeal by

the HOT theory to a state's being conscious only if one is aware of it ties the theory especially closely to our conception of what it is for a state to be conscious, again an advantage.

Dennett's puzzles about timing rest on there being nothing in subjective awareness or folk theory that could determine whether revision of a perceptual input occurs before or after it becomes conscious. The HOT theory offers a way to resolve that question. A perception becomes conscious only when there is a suitable HOT. And we can in principle determine whether the revision of content occurs before or after the onset of a HOT.

And we will see in sections V and VI that the HOT theory actually provides theoretical leverage for determining what types of mental states are likely to occur consciously in nonhuman animals, leading to a more nuanced picture than is available on Dennett's theory of consciousness.

As noted, Dennett sees merit in the HOT theory. But he also has some misgivings. One is a worry that the HOT theory tacitly involves what he calls the Cartesian Theater, on which there is a single location where all conscious states are observed. But that concern is unfounded. The HOT theory holds that a state is conscious in virtue of a HOT that makes one aware of being in that state. Each state is conscious in virtue of its own HOT; there is no single location, and no Cartesian Theater.

Dennett has also expressed doubts about applying the HOT theory to nonhuman animals, contending that animals that lack language could not have thoughts about what mental states they are in (2018, 169; 2022, 182-3). This is surprising, since Dennett himself denies that any states in nonhuman animals are conscious.

In any case, there is compelling evidence at least some nonhuman animals can have such thoughts. Great apes and some monkey species have psychological mechanisms that monitor and control their cognitive processes (Subias et al 2024). Such metacognitive processes must represent first-order mental states. And HOTs are cognitively less demanding than metacognition, since they involve no epistemic assessment. Such metacognitive ability strongly suggests that nonhuman animals have the cognitive capacity to form some HOTs.

Birch (2022) rejects higher-order (HO) theories generally because they rely on the human case; so he thinks they cannot help with nonhuman animals. But as Jacob Berger and Myrto Mylopoulos (2024) show, this is a mistake. HO theories all rely on the human case only to fix what it is to be a conscious state.

Other concerns have been raised about the HOT theory, and have been addressed elsewhere (e.g., Rosenthal 2005, 2022). But we can for now safely conclude that there is compelling reason to prefer that theory to Dennett's.

V. How Qualitative States Come to Be Conscious

On Dennett's FPO, a state comes to be conscious if it has the right effects. And the idea of multiple drafts competing to make it to consciousness suggests that winning that

competition is what results in a state's coming to have those effects, and so coming to be conscious. And it might well be that a draft typically wins that competition if it has greater utility for the organism than other contending drafts.

This utility would be due to the representational character of a draft, not to its being conscious. So it would be compatible with there being little if any added utility to a state's being conscious. And because winning the competition on Dennett's view is due to greater utility and also results in a state's being conscious, FPO would also fit with Dennett's claim, noted earlier, that we must understand consciousness in terms of some benefit it confers on the organism (Dennett 2022).

By contrast, a HO theory of consciousness holds that a state is conscious if an individual is subjectively aware of being in that state. So explaining why some states come to be conscious requires explaining how the relevant HO awareness arises. And if a state's being conscious adds little utility for the organism, added utility cannot explain how such HO awarenesses arise.

HO theories differ about the type of HO awareness they posit. On the higher-order-thought (HOT) theory that I have developed (e.g., 1993, 2005, 2022), that HO awareness consists of having a thought that one is in the state. Such HOTs are rarely conscious, and do not subjectively seem to rely on any inference. And as noted earlier, the usefulness of verbal report for indicating that a state is conscious points to a HOT theory, since a verbal report expresses a thought that one is in the relevant state.

Utility can explain why a draft would win its competition, since that utility would be due to the representational content of the draft. But if there is little added utility for a state to be conscious, utility cannot explain why a HOT arises, since the presence of a HOT constitutes a state's being conscious. So the HOT theory must explain that in some other way.

Following Dennett's methodological strictures, we can start by explaining how HOTs arise for humans, and then see how that explanation might apply to nonhuman animals. Doing so would support a reasonable inference, using the HOT theory, about whether HOTs are likely to arise in nonhuman animals, and so whether they are likely to be in any mental states that are conscious.

Adult humans readily recognize what visual inputs characteristically go with particular auditory or tactile inputs. Such cross-modal associations are automatic and natural for adult humans. But they are not given in the nature of things. So those associations must be learned, starting in early infancy (e.g., Held et al 2011). The infant must learn, for example, that the sight of mother's face goes with the sound of her voice, and the visual perception of a flat or curved surface goes with the corresponding tactile perceptions. And the infant must learn what bodily sensations, both pleasurable and painful, go with particular perceptible stimuli.

Learning these cross-modal associations is crucial to an infant's ability to interact with its environment. So there is considerable pressure for the infant to come to register which sensations from one modality go with those from another. And registering such associations requires representing the sensations that are associated.

There is no perceptual modality that could represent these cross-modal correlations. These associations can be represented only by thoughts that one type of sensation goes with another. The infant must have primitive thoughts whose contents represent these cross-modal associations.

Such thoughts represent oneself as being in states characterized by particular mental qualities. They are HOTs. The inescapable and pressing need to learn and represent cross-modal associations requires HOTs, which result in one's being aware of various mental states in respect of their qualitative character.

This explanation readily transfers to nonhuman animals. Nonhuman animals must also come to associate sensations of one modality with those of another. So they too must come to have HOTs about such sensations. The HOT theory explains how sensations come to be conscious both in humans and in very many nonhuman animals.

As noted earlier, Dennett doubts that nonhuman animals can form thoughts of the type this explanation posits, and others share such doubts. But they are unfounded. As Berger and Mylopoulos (2024, 98) stress, some nonhuman animals mentally represent conspecifics as perceiving things (e.g., Hare et al 2000). Such nonhuman animals have thoughts about perceptual states in respect of their representational properties. There is no serious reason to doubt that nonhuman animals can form thoughts of the type this explanation calls for. And the HOT theory to one side, many nonhuman animals plainly do represent cross-modal associations. It is unclear how else they could do so.

There is a second concern about the foregoing explanation. One might insist that the cross-modal associations that one must represent do not hold between qualitative states at all, but only between perceptible properties that are accessible by different modalities. And then the thoughts that represent those associations would not be HOTs after all, and so not relevant to how qualitative states come to be conscious.

But what is needed is a representation of how perceptible properties accessible by distinct modalities appear to one in perception. The infant or nonhuman animal must respond to how inputs from one modality correlate with inputs from another. The infant or nonhuman animal has no access to the objective perceptible properties apart from the sensations that result. The mental representation of the relevant correlations are HOTs.

One might also object to the foregoing explanation by contending that the relevant cross-modal associations are simply the product of statistical learning, which might not give rise to HOTs that such associations hold. But many cross-modal associations will be represented in thought; an infant's expectation to see its mother on hearing her voice will be represented conceptually. Statistical learning doubtless plays a significant role, but very many cross-modal associations will also be represented conceptually. The infant doubtless represents its mother conceptually, and so must conceptually register relevant visual and auditory inputs.

Cross-modal calibration has great utility, which explains why it occurs. And learning such calibrations results in HOTs about the qualitative character of perceptual states.

But the utility results only from the representation of cross-modal associations, not from the attendant consciousness of perceptual states. So these considerations do not run afoul of the observation that a state's being conscious adds little if any utility to the state.

VI. How Thoughts Come to Be Conscious

The foregoing explanation will not apply to the occurrence in humans of HOTs about mental states in respect of their conceptual content. But many mental states in humans are conscious in respect of those conceptual properties. So the HOT theory needs a different explanation for those cases.

And there is an alternative explanation, which relies instead on a somewhat elaborate human linguistic ability. Human children initially learn to conceptualize others as having thoughts by hearing adults describe themselves and others in that way. As the child catches on, it begins to ascribe thoughts to others on the basis of behavior and context (Rosenthal 2025b; Sellars 1956).

As adults ascribe thoughts to the child, the child also learns to ascribe thoughts to itself. Sometimes this self-ascription may still rely on behavior and context. But such self-ascription will in time come to be independent of those factors. The child will describe itself as having a particular thought simply because it subjectively seems to the child that it does, independently of any observations.

The primary way in which this comes to happen relies on the child's being disposed to say particular things. The child learns that when one is disposed to say something, it is appropriate to describe oneself as thinking that thing. The content of what one is disposed to say goes with the content of a thought that would be expressed by that speech act.

And the child will often be aware of being disposed to say things, much as one is aware of being disposed to do things in general, such as getting up or going someplace. So mastering the association of one's being disposed to say something with one's having the corresponding thought results in the child's coming to be aware of many of its thoughts. This awareness of having a thought is verbally expressed by saying that one has a particular thought. So the awareness is constituted by the child's having a HOT that it has the thought in question.

Being aware of a thought initially always follows an awareness of being disposed to say something. But in time that association becomes second nature, so that an awareness of having a thought results from being disposed to say something even when one is not aware of being so disposed. And then the awareness of having a thought will subjectively seem to occur without any psychological antecedent.

From a theoretical point of view, the HOT that one thinks something is then subjectively unmediated. These considerations explain the compelling sense we have that there is some important connection between consciousness and speech, a connection that it is

often noted but goes unexplained. (For similar considerations see Rosenthal 2005, ch. 10, and Rosenthal 2025b, §VII.)

It is unlikely that any other process could give rise to HOTs that one has particular thoughts. We do not think of nonverbal behavior as having content that could match a thought expressed by that nonverbal behavior. So it is unlikely that any such HOTs occur in creatures that lack the linguistic abilities that figure in this explanation. On a HOT theory, conceptual mental states will not occur consciously in nonhuman animals, nor in prelinguistic human infants.

This conclusion is arguably compatible with our pretheoretic sense both about infants and about nonhuman animals. We have a strong sense that many perceptual states of nonhuman animals and human infants are conscious. But we also tend to be agnostic about whether their thoughts and other conceptual states are conscious. It is a distinct advantage of the HOT theory over Dennett's FPO that the HOT theory explains this difference in our pretheoretic convictions, and also enables predictions about nonhuman animals in respect of conceptual states as against sensations.

VII. Experimental Tests

A verbal report that one is in some state indicates that one is aware of being in that state, and so indicates that it is conscious, but verbal report is unavailable for nonhuman animals. Button presses and other substitutes for verbal report are also intended to indicate when a state occurs consciously. But Dennett rightly notes that such nonverbal substitutes might simply be conditioned responses to unconscious states.

A direct test that states in nonhuman animals are conscious may not be forthcoming given our current states of knowledge. But that does not rule out an indirect approach for conscious states. One can design a test that directly indicates that a mental state fails to occur consciously, and then find that this test does not yield that result for other mental states. We could then conclude that those other mental states are conscious.

Here is an example. Alan Cowey and Petra Stoerig (1995) tested three macaque monkeys whose left striate cortex had been surgically ablated, resulting in blindsight in the right half of their visual field. When stimuli were presented to their intact field, the monkeys' task was to touch a location on a screen that indicated the position of the stimulus. But sometimes there was also a screen position they could touch to indicate that no stimulus had been presented. When a stimulus was presented to the monkeys' hemianopic field, they would touch that blank screen position whenever it was available. But when the blank option was not available, the monkeys would instead often correctly indicate the position of the stimulus.

As Cowey and Stoerig note (1995, p. 249), these results strongly resemble the behavior of human blindsight patients, who deny seeing stimuli presented to their blind field but when pressed to guess are largely correct about the properties of the stimulus. When the blank option is unavailable for the monkeys, there is some psychological pressure,

as with forced-choice guessing, to classify the stimulus, and the monkeys mainly guess correctly. But they prefer the blank option whenever it is available.

The monkeys were trained to use the blank option to indicate that nothing was presented. So choosing that option when it is available shows that they did not consciously see anything. But the monkeys never use the blank option when stimuli are presented to their intact field. So the test for unconscious seeing fails when stimuli are presented to the intact field. We can infer that seeing in that case is conscious.

Varying whether the blank option was available allowed Cowey and Stoerig to demonstrate that presentations made to the hemianoptic field were seen unconsciously, in contrast with presentations to the intact field. And that enabled an experimental way to distinguish seeing that is conscious from seeing that is not.

Training the monkeys doubtless involved reinforcement conditioning. But Dennett's concern about reinforcement conditioning is that the resulting behavior might just be a conditioned response to an unconscious state. Because the test here is whether a state is unconscious, Dennett's concern does not apply.

This strategy, which relies on directly testing for states to be unconscious, relies on what we know about human blindsight. So it reflects Dennett's methodological insistence on anchoring our investigation into nonhuman animals by the case of humans.

Change blindness offers another way to implement this experimental strategy. Change blindness occurs when an individual is unaware of a reasonably salient change in the individual's visual field (e.g., Simons & Levin, 1997; Grimes 1996). But even when human participants are unaware of such a change, they nonetheless often perceive that change unconsciously (Fernandez-Duque & Thornton 2000, Thornton & Fernandez-Duque 2001; replicated by Laloyaux et al 2006).

Change blindness has been demonstrated in nonhuman animals as varied as chimpanzees and pigeons (Herbranson 2022). And measures of change detection have also been studied for nonhuman animals (Wright 2022). So experiments could be designed to show that even when a nonhuman animal indicates that it does not see a change, it nonetheless sometimes perceptually registers that change.

If a nonhuman animal indicates it sees no change but nonetheless registers that change perceptually, such perceptual registration is unconscious. And those cases would contrast with cases in which the animal indicates that it does see a change. Since in one case the animal indicates that it does not see anything and in the other that it does, we can infer that seeing in the second case is conscious, in contrast with seeing in the first case. Like Cowey and Stoerig's blindsight strategy, this procedure is independent of any worry about reinforcement conditioning, and also rests on our knowledge of change blindness in the human case. This proposal is well worth exploring.

Dennett does not consider such empirical strategies, nor to my knowledge do any of the current literature on consciousness in nonhuman animals. This is unfortunate, but it is readily explained. Mental states are always conscious, according to Dennett, and the subpersonal drafts that fail to be conscious are states of a different type. And that

prevents showing that a state is conscious because a test for its being unconscious fails. Dennett's methodological injunctions are sound and important. But his theory of consciousness restricts him to analogues in nonhuman animals of verbal report and the equivalent, which he reasonably finds wanting.

On the HOT theory, by contrast, whether a state is conscious or unconscious depends on whether or not a HOT occurs, and the state itself is of the same type in both cases. Since the difference between conscious states and mental states that are not conscious is central to the theory, it is natural on the theory to determine that a state is conscious by showing that a test for mental states to be unconscious fails.

Because much of the current literature relies on behavioral markers, that literature is also limited in the type of test it can consider. Behavioral markers are intended to indicate the occurrence of conscious states. So being wedded to behavioral markers focuses attention exclusively on testing for a state to be conscious, as against its being unconscious. And reliance on markers will likely distract from the search for tests that are more sound (cf. Lau 2024).

The failure to take unconscious mental states seriously restricts the kind of empirical tests one will consider. But more generally it also distorts our conception of what consciousness itself is. We can best understand what it is for a mental state to be conscious by the way conscious states differ from mental states that fail to be conscious. If one does not have a theoretical way to address that contrast, one's thinking about conscious states will likely either rely solely on first-person access, as with Nagel, or equate consciousness with mentality, as with Dennett. Both pictures distort how we think about consciousness.

The implications of the HOT theory rehearsed in section V suggest that some perceptual states in many nonhuman animals are likely to occur consciously. And despite Dennett's concerns, experimental tests are available to establish that in at least some cases.

I have argued that some mental states in many nonhuman animals very likely do occur consciously. Still, it is worth stressing in closing that the ethical treatment of nonhuman animals should not hinge on whether that is so. We sensibly take ethical concerns to arise initially for humans. So it may be tempting to hold that we should extend ethical consideration to nonhuman animals only if they are relevantly similar to humans. And one might regard nonhuman animals' being in conscious mental states as a salient similarity for those purposes.

Conscious pain and suffering are doubtless worse than pain and suffering that might occur without being conscious. But even unconscious pain and suffering would be significantly bad for an organism. If those states occurred unconsciously, they would disrupt an organism's functioning and impair its quality of life and well-being. And the organism would doubtless strongly desire, even if unconsciously, that those conditions cease. So pain and suffering arguably deserve ethical consideration whether or not they occur consciously. More generally, ethical consideration is warranted for nonhuman animals independently of whether any of their mental states are conscious.

References

Allen, Colin, and Michael Trestman (2016), "Animal Consciousness," *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta & Uri Nodelman, https://plato.stanford.edu/archives/win2020/entries/consciousness-animal/

Andrews, Kristin (2016), "A Role for Folk Psychology in Animal Cognition Research," in Animals: New Essays, ed. Andreas Blank, Munich: Philosophia, pp. 205-226.

Andrews, Kristin (2020), The Animal Mind: An Introduction to the Philosophy of Animal Cognition, 2nd edition, London: Routledge.

Andrews, Kristin (2024), "'All Animals Are Conscious': Shifting the Null Hypothesis in Consciousness Science," Mind & Language, 39, 3 (June): 415-433.

Andrews, Kristin, Jonathan Birch, and Jeff Sebo (2025), "Evaluating Animal Consciousness," Science, 387, 6736 (February 21): 822-824.

Bayne, Tim, and Nicholas Shea (2020), "Consciousness, Concepts, and Natural Kinds," Philosophical Topics, 48, 1 (Spring): 65-83.

Berger, Jacob, and Myrto Mylopoulos (2024), "HOTT and Heavy: Higher-Order Thought Theory and the Theory-Heavy Approach to Animal Consciousness," Synthese, 203, 3 (March): article 98.

Birch, Jonathan (2022), "The Search for Invertebrate Consciousness,"Noûs, 56, 1 (March): 133-153.

Birch, Jonathan (2024), The Edge of Sentience, Oxford: Oxford University Press.

Birch, Jonathan, Alexandra K. Schnell, and Nicola S. Clayton, (2020), "Dimensions of Animal Consciousness," Trends in Cognitive Sciences, 24, 10 (October): 789–801.

Birch, Jonathan, Donald M. Broom, Heather Browning, Andrew Crump, Simona Ginsburg, Marta Halina, David Harrison, Eva Jablonka, Andrew Y. Lee, François Kammerer, Colin Klein, Victor Lamme, Matthias Michel, Françoise Wemelsfelder, and Oryan Zacks (2022), "How Should We Study Animal Consciousness Scientifically?", Journal of Consciousness Studies, 29, 3-4: 8-28.

Bourne, James A., and Maria Concetta Morrone (2017), "Plasticity of Visual Pathways and Function in the Developing Brain: Is the Pulvinar a Crucial Player?", Frontiers in Systems Neuroscience, 11 (February): Article 3.

Cohen, Michael A., and Daniel C. Dennett (2011), "Consciousness Cannot Be Separated from Function," Trends in Cognitive Sciences, 15, 8 (August): 358-364.

Cohen, Michael A., and Daniel C. Dennett (2012), "Response to Fahrenfort and Lamme: Defining Reportability, Accessibility and Sufficiency in Conscious Awareness," Trends in Cognitive Sciences, 16, 3 (March): 139-140.

Cowey, Alan (2010), "The Blindsight Saga," Experimental Brain Research, 200, 1 (January): 3-24.

Cowey, Alan, and Petra Stoerig (1995), "Blindsight in Monkeys," Nature, 373, 6511 (January 19): 247-249.

Danckert, James, and Yves Rossetti (2005), "Blindsight in Action: What Can the Different Sub-types of Blindsight Tell Us about the Control of Visually Guided Actions?", Neuroscience and Biobehavioral Reviews, 29, 7: 1035-1046.

de Gelder, Beatrice, Marco Tamietto, Geert van Boxtel, Rainer Goebel, Arash Sahraie, Jan van den Stock, Bernard M. C. Stienen, Lawrence Weiskrantz, and Alan Pegna (2008), "Intact Navigation Skills after Bilateral Loss of Striate Cortex," Current Biology, 18, 24 (December 23): r1128-r1129.

de Gelder, Beatrice, Nicholas Humphrey, and Alan J. Pegna (2024), "On the Bright Side of Blindsight: Considerations from New Observations of Awareness in a Blindsight Patient," Cerebral Cortex, bhae456.

Dennett, Daniel C. (1969), Content and Consciousness, New York: Routledge & Kegan Paul.

Dennett, Daniel C. (1971), "Intentional Systems," The Journal of Philosophy, 68, 4 (February 26): 87-106.

Dennett, Daniel C. (1989), The Intentional Stance, Cambridge, Massachusetts: MIT Press.

Dennett, Daniel C. (1991), Consciousness Explained, Boston: Little, Brown and Company.

Dennett, Daniel C. (1995), "Animal Consciousness: What Matters and Why," Social Research, 62, 3 (Fall): 691–710.

Dennett, Daniel C. (2018), "Reflections on David Rosenthal," in The Philosophy of Daniel Dennett, ed. Bryce Huebner, New York: Oxford University Press, pp. 165-170.

Dennett, Daniel C. (2022), "A Challenge: Can Self-Monitoring Be Achieved without Consciousness?", in Qualitative Consciousness: Themes from the Philosophy of David Rosenthal, ed. Josh Weisberg, Cambridge: Cambridge University Press, pp. 181-186.

Dennett, Daniel C., and Marcel Kinsbourne (1992), "Time and the Observer: The Where and When of Consciousness in the Brain," Behavioral and Brain Sciences, 15, 2 (June): 183-247.

Dienes, Zoltan (2008), "Subjective Measures of Unconscious Knowledge," in Models of Brain and Mind: Physical, Computational and Psychological Approaches, ed. Rahul Banerjee and Bikas K. Chakrabarti, Amsterdam: Elsevier, pp. 49-64.

Dijksterhuis, Ap, Maarten W. Bos, Loran F. Nordgren, and Rick B. van Baaren (2006), "On Making the Right Choice: The Deliberation-without-attention Effect," Science, 311, 5763 (February 17): 1005–1007.

Fernandez-Duque, Diego, and Ian M. Thornton (2000), “Change Detection without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System?”, Visual Cognition, 7, 1-2-3 (January–March): 324–344.

Fleming, Stephen M., and Matthias Michel (2024), “Sensory Horizons and the Functions of Conscious Vision,” PsyArXiv, (November 22): https://doi.org/10.31234/osf.io/4tqxs

Ginsburg, Simona, and Eva Jablonka (2019), The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness. Cambridge, Massachusetts: MIT Press.

Godfrey-Smith, Peter (2020), Metazoa: Animal Life and the Birth of the Mind, New York: Farrar, Straus, and Giroux.

Gottlieb, Joseph (2024), “Unlimited Associative Learning and the Theory-Light Approach to Non-human Consciousness,” Journal of Consciousness Studies, 31 11-12 (December): 85-109.

Grimes, John (1996), “On the Failure to Detect Changes in Scenes across Saccades,” Perception, ed. Kathleen Akins, New York: Oxford University Press, pp. 89‑110.

Hare, Brian, Josep Call, Bryan Agnetta, and Michael Tomasello (2000), “Chimpanzees Know What Conspecifics Do and Do Not See,” Animal Behaviour, 59, 4 (April): 771-785/

Helfrich-Förster, Charlotte (2018), “Sleep in Insects,” Annual Review of Entomology, 63 (January): 59-86.

Held, Richard, Yuri Ostrovsky, Beatrice de Gelder, Tapan Gandhi, Suma Ganesh, Umang Mathur, & Pawan Sinha (2011), “The Newly Sighted Fail to Match Seen with Felt”, in Nature Neuroscience, 14, 5 (May): 551–553.

Herbranson, Walter T. (2022), “Change Blindness,” in Encyclopedia of Animal Cognition and Behavior, ed. Jennifer Vonk and Todd K. Shackelford, Cham, Switzerland: Springer, pp. 1279-1282.

Humphrey, Nicholas (1972), “Helen: ‘A Blind Monkey Who Sees Everything,’ A Case of Super-Blindsight,” YouTube, 1 minute, 28 seconds, https://www.youtube.com/watch?v=PQl-cT_G5x4

Humphrey, Nicholas (1974), “Vision in a Monkey without Striate Cortex: A Case Study,” Perception, 3, 3 (September): 241-255.

Humphrey, Nicholas (2002), “The Uses of Consciousness,” in Humphrey, The Mind Made Flesh: Essays from the Frontiers of Evolution and Psychology, Oxford: Oxford University Press, pp. 65–85.

Humphrey, Nicholas, and Lawrence Weiskrantz (1967), “Vision in Monkeys after Removal of the Striate Cortex,” Nature, 215, 5101 (August 5): 595-597.

Kolers. Paul A., and Michael von Grünau (1976), "Shape and Color In Apparent Motion," Vision Research. 16, 4: 329—335.

Laloyaux, Cédric, Arnaud Destrebecqz, and Axel Cleeremans (2006), "Implicit Change Identification: A Replication of Fernandez-Duque and Thornton (2003)," Journal of Experimental Psychology: Human Perception and Performance, 32, 6 (December): 1366-1379.

Lau, Hakwan (2024), "Premature Declarations on Animal Consciousness Hinder Progress," The Transmitter (November 18): https://doi.org/10.53053/BCOU1212

Maćkiewicz, Bartosz, Katarzyna Kuś, and Witold M. Hensel (2023), "The Influence of Philosophical Training on the Evaluation of Philosophical Cases: A Controlled Longitudinal Study," Synthese, 202, 4 (October): Article 113.
Marcel, Anthony J. (1983), "Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes," Cognitive Psychology, 15, 2 (April): 238–300.

Michel, Matthias (manuscript), "On Confusions about the Functions of Consciousness."

Nagel, Thomas (1974), "What Is It Like to Be a Bat?", The Philosophical Review, 83, 4 (October): 435–50; reprinted in Nagel, What Is It Like to Be a Bat?, Oxford: Oxford University Press, 2024, pp. 1-29 and 56-64.

Öğmen, Haluk, and Bruno G. Breitmeyer (2006), eds., The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual rocesses, Cambridge, Massachusetts: MIT Press.

Peirce, Charles Sanders, and Joseph Jastrow (1884), "On Small Differences in Sensation," Memoirs of the National Academy of Sciences, 3: 73–83.

Phillips, Ian (2018), "Unconscious Perception Reconsidered," Analytic Philosophy, 59, 4 (December): 471–514.

Robinson, Zack, Corey J. Maley, and Gualtiero Piccinini (2015), "Is Consciousness a Spandrel?", Journal of the American Philosophical Association, 1, 2 (Summer): 365-383.

Rosenthal, David (1993), "State Consciousness and Transitive Consciousness," Consciousness and Cognition, 2, 4 (December 1993): 355-363.

Rosenthal, David (2005), Consciousness and Mind, Oxford: Clarendon Press.

Rosenthal, David (2008), "Consciousness and Its Function," Neuropsychologia, 46, 3: 829-840.

Rosenthal, David (2010), "How to Think about Mental Qualities," Philosophical Issues, 20 (October): 368-393.

Rosenthal, David (2012), "Awareness and Identification of Self," in Consciousness and the Self: New Essays, ed. JeeLoo Liu and John Perry, Cambridge: Cambridge University Press, pp. 22-50.

Rosenthal, David (2018), "Seeming to Seem," in The Philosophy of Daniel Dennett, ed. Bryce Huebner, New York: Oxford University Press, pp. 133-164.

Rosenthal, David (2022), "Mental Appearance and Mental Reality," in Qualitative Consciousness: Themes from the Philosophy of David Rosenthal, ed. Josh Weisberg, Cambridge: Cambridge University Press, pp. 243-271.

Rosenthal, David (2025a), "Objective Foundations for the Study of Mental Qualities," Philosophy and the Mind Sciences, VI.

Rosenthal, David (2025b), "Thought, Consciousness, and the Given," European Journal of Philosophy, DOI: 10.1111/ejop.13039

Sellars, Wilfrid (1956), "Empiricism and the Philosophy of Mind," in Herbert Feigl and Michael Scriven (eds.), Minnesota Studies in the Philosophy of Science, vol. I. Minneapolis: University of Minnesota Press, pp. 253-329.

Simons, Daniel J., and Daniel T. Levin (1997), "Change Blindness," Trends in Cognitive Sciences, 1, 7 (October): 261-267.

Sklar, Asael Y., Nir Levy, Ariel Goldstein, and Ran Hassin (2012), "Reading and Doing Arithmetic Nonconsciously," Proceedings of the National Academy of Sciences, 109, 48 (November 27): 19614–19619.

Stoerig, Petra (2011), "Task-rrelevant Blindsight and the Impact of Invisible Stimuli," Frontiers in Psychology, 2 (April): Article 66.

Subias, Lorraine, Subias Noriko Katsu, and Kazunori Yamada (2024), "Metacognition in Nonhuman Primates: A Review of Current Knowledge," Primates, https://doi.org/10.1007/s10329-024-01169-x

Thornton, Ian M., and Diego Fernandez-Duque (2001), "An Implicit Measure of Undetected Change," Spatial Vision, 14, 1 (January): 21-44.

Timmermans, Bert, and Alex Cleeremans (2015), "How Can We Measure Awareness? An Overview of Current Methods," in Behavioral Methods in Consciousness Research, ed. Morten Overgaard, Oxford: Oxford University Press.

Veit, Walter (2023), A Philosophy for the Science of Animal Consciousness, New York: Routledge.

Waroquier, Laurent, David Marchiori, Olivier Klein, and Axel Cleeremans (2010), "Is It Better to Think Unconsciously or to Trust Your First Impression? A Reassessment of Unconscious Thought Theory," Social Psychological and Personality Science, 1, 2 (April): 111-118.

Weiskrantz, Lawrence (1990), Blindsight: A Case Study and Implications, Oxford: Clarendon Press.

Weiskrantz, Lawrence (1997), Consciousness Lost and Found, Oxford: Oxford University Press.

Wright, Anthony A. (2022), “Change Detection,” in Encyclopedia of Animal Cognition and Behavior, ed. Jennifer Vonk and Todd K. Shackelford, Cham, Switzerland: Springer, pp. 1282-1292.